Exhibit 4

Textbook chapter from
Recombinant DNA, 2nd edition,
Scientific American Books, New York (1992)

Recombinant DNA

SECOND EDITION

James D. Watson
COLD SPRING HARBOR LABORATORY

Michael Gilman
COLD SPRING HARBOR LABORATORY

Jan Witkowski
BANBURY CENTER, COLD SPRING HARBOR LABORATORY

Mark Zoller
GENENTECH, INC.

SCIENTIFIC AMERICAN BOOKS

Distributed by
W. H. Freeman and Company
New York

foreign protein is produced as an in-frame fusion with a bacterial protein. Often fusion proteins are more soluble and therefore easier to work with, and the fused polypeptide sometimes provides a tag or an activity that facilitates purification of the protein from cell extracts. A protein can be targeted for secretion by fusing the coding sequence to a segment of a bacterial protein containing signals for secretion out of the cell. Protein produced in *E. coli* can be used for functional studies if it retains its biochemical activity or as a source of antigen for immunizing animals for the production of specific antisera against the protein. Using similar strategies, cloned cDNAs can be expressed in mammalian cells as well. Mammalian proteins are more likely to be functional when expressed in mammalian cells. For example, expression of a cDNA encoding a protein believed to be a growth factor receptor should result in the expression of binding sites for the growth factor on the surface of the expressing cell, as discussed above for the EPO receptor (Figure 7-12).

Genomic Fragments Are Cloned in Bacteriophage

Cloning of chromosomal genes is relatively easy once the corresponding cDNAs have been isolated and are available as probes. It is essential to analyze genomic DNA if regulatory sequences outside the coding sequences of a gene are to be studied. As we shall see in Chapter 8, analysis of chromosomal genes revealed unforeseen complexities in the organization of eukaryotic genes. Sometimes the abundance of a particular mRNA is so low that a cDNA derived from it is difficult to find in a cDNA library using the available probes. Instead, a genomic clone is first isolated, which serves as a better probe to find the rare cDNA clone in the library. Although genomic clones usually contain noncoding sequences (see Chapter 8), a fragment from the genomic clone containing coding sequences can be isolated and used to screen the library. In addition, the genomic DNA can be introduced into a cell line and expressed using a strong viral promoter. Often under a different promoter, the mRNA is highly expressed, and then its cDNA can be isolated from this cell line.

Early attempts to prepare genomic DNA libraries in plasmids were not successful. While cDNAs are relatively small and of an appropriate size for cloning in plasmids, it soon became clear that plasmids were not suitable for cloning large segments of chromosomal DNA. Because small plasmids replicate more efficiently than large ones, the latter are selected against and portions of the cloned DNA are progressively lost. In contrast, large chromosomal DNA fragments (of about 15,000 bp) are stable when inserted into the DNA of special strains of λ phage. Already at the 1975 Asilomar conference it had been suggested that λ phage could be mutated so that it would be unable to insert its DNA into that of host *E. coli* cells, and thus it would be at least as safe as, if not safer than, disabled plasmid vectors. Such λ vectors exploited the fact that the entire central section of λ phage DNA is not necessary for its replication in *E. coli*, but functions only to ensure the integration of the phage DNA into the host bacterial chromosome during its lysogenic phase. Strains of λ have been created in which recognition sites for a restriction enzyme are located so as to leave intact the left and right end fragments (the *arms*) of the phage DNA that are essential for its replication and packaging (Figure 7-20). After digestion with the enzyme, these end fragments, because of their relatively large sizes, can easily be isolated and can be used to make new λ-like phages containing one left arm, one right arm, and a foreign DNA insert. Maturation of λ phage requires that its DNA chromosome be approximately 45 kb long; thus the only DNAs constructed in vitro that can multiply following such manipulations are recombinant molecules that consist of DNA inserts about 15 kb in length flanked by phage arms.

Genomic libraries were first constructed by cutting genomic DNA to completion with a restriction enzyme such as *Eco*RI and cloning these fragments into a λ vector. But many genomic *Eco*RI fragments are either too large or too small to yield viable phages, and therefore genes carried on such fragments would be missing from a library made this way. Instead, to make a library in which all genomic sequences are represented, genomic DNA is broken in as nearly random a fashion as possible (Figure 7-20). Random fragments of a size optimal for packaging in λ vectors are selected and cloned. Typically several million independent λ clones carrying genomic fragments are

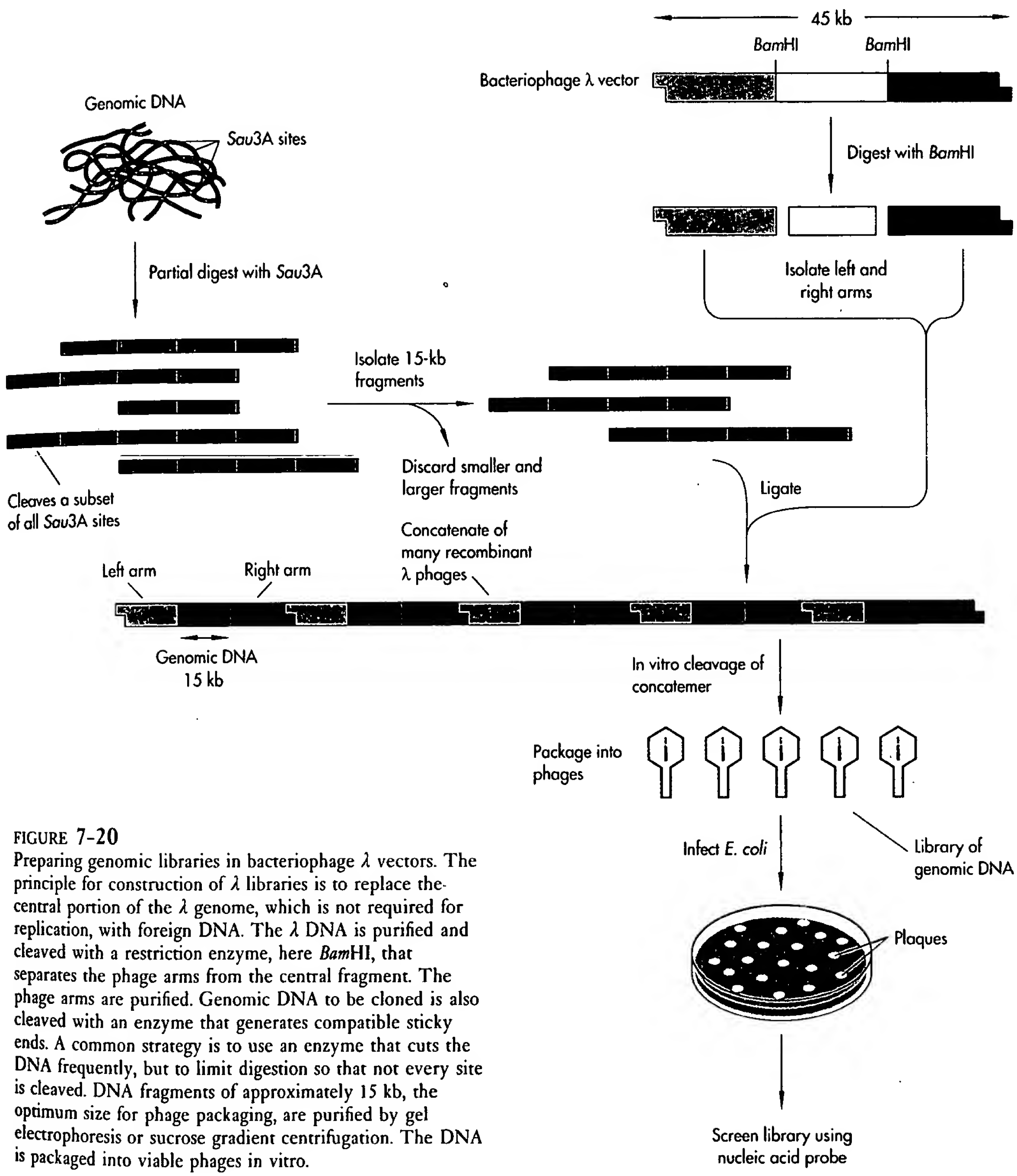


FIGURE 7-20
Preparing genomic libraries in bacteriophage λ vectors. The principle for construction of λ libraries is to replace the central portion of the λ genome, which is not required for replication, with foreign DNA. The λ DNA is purified and cleaved with a restriction enzyme, here *Bam*HI, that separates the phage arms from the central fragment. The phage arms are purified. Genomic DNA to be cloned is also cleaved with an enzyme that generates compatible sticky ends. A common strategy is to use an enzyme that cuts the DNA frequently, but to limit digestion so that not every site is cleaved. DNA fragments of approximately 15 kb, the optimum size for phage packaging, are purified by gel electrophoresis or sucrose gradient centrifugation. The DNA is packaged into viable phages in vitro.

obtained. Mammalian genomes contain about 3×10^9 bp of DNA. If the average insert size is 15,000 bp, approximately 200,000 phages will carry a genome's worth of DNA. Because, however, the phages carry random DNA fragments, the first 200,000 phages selected will carry some sequences more than once and others not at all. To ensure that all sequences in the genome are present at least once, simple statistical calculations show that roughly 1 to 2 million phages must be screened (to understand why, simply imagine how many playing cards you would need to draw from a deck to ensure that you have at least one card of each suit; unless you are very lucky, you will need to draw more than four).

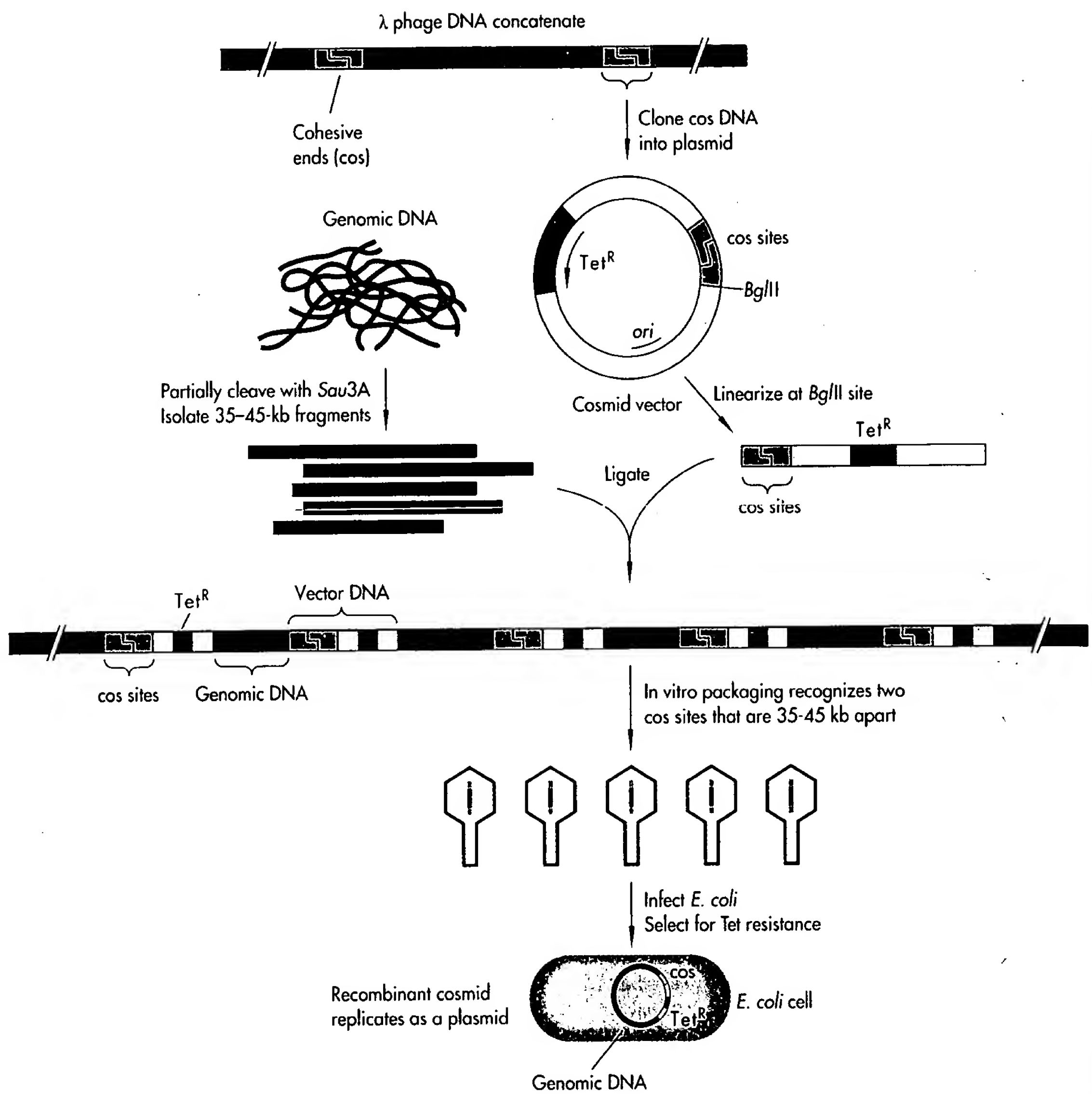


FIGURE 7-21
Cosmids. Cosmids are plasmid vectors that carry the cos sites from the λ phage as well as a standard plasmid origin of replication and a drug-resistance gene (Tet^R). To clone genomic DNA into a cosmid vector, the vector is linearized with a restriction enzyme, here *Bgl*II, and the genomic DNA is partially digested with *Sau*3A, which leaves *Bgl*II-compatible ends. DNA fragments in the 35 to 40-kb range are isolated and ligated to linearized vector DNA, forming tandem arrays of vector and genomic DNA fragments (not drawn to scale). A λ packaging extract recognizes and packages any ligated DNA that carries two cos sites 35 to 45 kb apart. These segments of DNA are introduced into *E. coli* by infection and replicate as drug-resistance plasmids.

Cosmids Allow the Cloning of Large Segments of Genomic DNA

The size of eukaryotic DNA segments that can be carried in λ phages is limited. It is clear from genomic analysis of a variety of organisms that many genes are larger than this, some as large as 1000 kb. Genes of this size have to be cloned as a set of overlapping genomic fragments. *Cosmid* vectors are hybrids, derived from plasmids and λ phages, that facilitate genomic cloning by being able to carry approximately 45 kb of foreign DNA, three times more than that of phage vectors (Figure 7-21).

The λ phage contains at each end single-stranded complementary stretches of DNA, the so-called *cos* sites. During the normal life cycle of λ, hundreds of copies of newly replicated phage DNA form long chains, or *concatamers,* each λ genome being joined to the next one in the chain through the cos sites. The λ packaging enzymes chop this concatamer into λ-sized pieces by recognizing two cos sites approximately 45 kb apart, cleaving this unit, and packaging it into phages. Thus, the cos sites are all that is necessary for packaging DNA into phage and cosmid vectors. To make a cosmid library, eukaryotic DNA is cleaved with a restriction enzyme under conditions that yield relatively large pieces of DNA. This DNA is then ligated to the cosmid, which has been cleaved with a restriction enzyme that leaves ends complementary to the cleaved genomic DNA. The ligated DNA is packaged in vitro into phages and introduced into *E. coli* by infection. Once inside the *E. coli* cell, the cosmid replicates and can be recovered from the cell as a plasmid. Cosmids represent an important link between λ vectors, which contain up to 15 kb of DNA, and YAC vectors (Chapter 29), which carry over 100 kb of DNA.

Chromosome Walking Is Used to Analyze Long Stretches of Eukaryotic DNA

The isolation of genomic DNA and the analysis of genomic organization must be done in a systematic way. An efficient way of doing this is to use one recombinant phage or cosmid to isolate another that contains overlapping information from the genome. This technique, known as *chromosome walking,* depends on obtaining a small segment of DNA from one end of the first recombinant and using this piece of DNA to rescreen the library to obtain additional recombinants containing that piece of DNA and the next portion of the genome (Figure 7-22). The second recombinant is used to obtain a third, and so on, to yield a set of overlapping cloned segments. Of course, the small piece of DNA used to rescreen the library must be a single-copy element in the genome; if it is a repeated sequence, many unrelated recombinants will be identified. It may be difficult to clone if it is unstable or toxic in *E. coli.* These situations are molecular roadblocks that may take particular effort to circumvent. Chromosome walking is the only way to search for a gene when its position on a chromosome is only approximately known. We will see in Chapter 26 how chromosome walking is used to clone human disease genes.

Southern and Northern Blotting Procedures Analyze DNA and RNA

Once a cloned cDNA is isolated, it can be used to analyze gene expression and the organization of the genomic DNA. The most widely used method is called *blotting* (Figure 7-23). *Southern blotting,* developed by E. M. Southern, is an extremely powerful tool for analyzing gene structure. To do a Southern blot, genomic DNA is cut with one or several restriction enzymes and the resultant fragments are separated by size on an agarose gel. The gel is then overlaid with a sheet of nitrocellulose filter or nylon membrane, and a flow of buffer is set up through the gel toward the nitrocellulose filter. This causes the DNA fragments to be carried out of the gel onto the filter, where they bind. Thus a "replica" of the DNA in the gel is created on the nitrocellulose. A labeled probe, specific for the gene under study, is then hybridized to the DNA molecules bound on the filter. This probe can be a purified RNA, a cloned cDNA, or a short synthetic oligonucleotide. The labeled probe will hybridize to the specific molecules containing a complementary sequence. Autoradiography of the nitrocellulose filter will result in a pattern of bands indicating the number

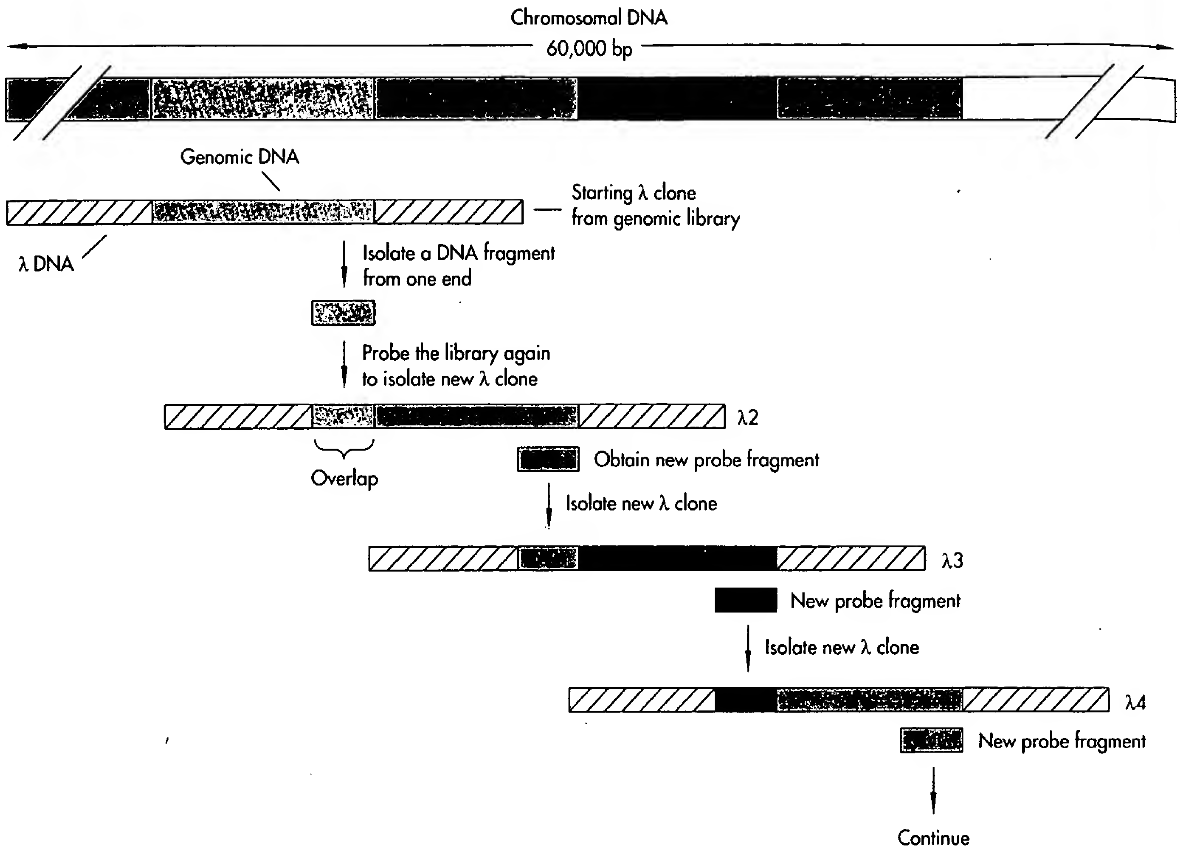
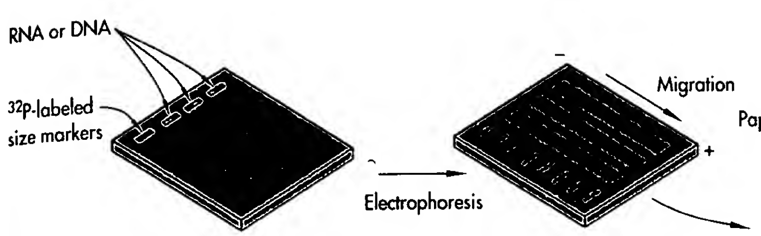


FIGURE 7-22
Chromosome walking. This method is used to move systematically along a chromosome from a known location. A cloned probe is used to isolate phage clones carrying genomic fragments from that region of the chromosome. A small DNA fragment from the end of the largest phage clone is used to rescreen the same library. Among the clones recovered are new phages (for example, $\lambda 2$) that carry the probe sequence but whose sequences also extend farther along the chromosome. A new probe is generated from the far end of one of these phages and used to screen the library again to isolate new clones extending still farther. Hundreds of kilobases of contiguous chromosomal DNA can be cloned by repeated cycles of walking. The colored regions in the figure indicate genes, and the hatched regions are λ DNA.

and size of the DNA fragments complementary to the probe. A physical map of the gene, consisting of the positions of landmark restriction sites, can be produced from the sizes of the fragments. Such a *restriction map* can be used to compare the DNA sample with others, allowing, for example, the detection of deletions or other rearrangements involving the gene.

Southern blotting is useful for detecting major gene rearrangements and deletions found in a variety of human diseases (see Chapters 18 and 27). It can also be used to identify structurally related genes in the same species and homologous genes in other species. Southern blots to a panel of genomic DNAs from a collection of organisms, *zoo blots*, reveal the degree of evolutionary conservation of a gene. For example, Southern blotting identified genes in yeast related to the *RAS* oncogene in human tumors, a remarkable example of evolutionary conservation. In fact, as we will see in Chapter 13, human *RAS* genes are functional in yeast.



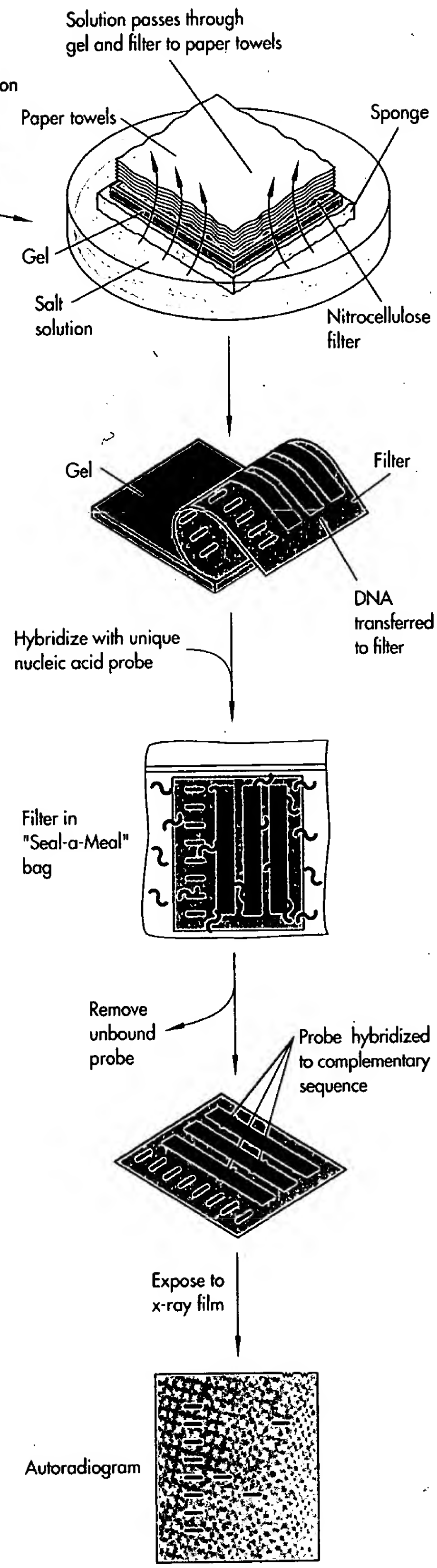


FIGURE 7-23
Analyzing DNA and RNA by gel electrophoresis and blotting. DNA cleaved with restriction enzymes or RNA isolated from cells is applied to an agarose gel and electrophoretically separated by size. The nucleic acids in the gel are transferred to a nitrocellulose filter to make a precise replica of the gel. This is usually done by placing the gel atop a sponge sitting in a tray of buffer. The filter is laid over the gel and covered with a stack of paper towels that acts as a wick, pulling buffer up through the sponge, gel, and filter. DNA or RNA fragments from the gel are carried up onto the filter, where they stick tightly. The filter is removed and hybridized with a radioactively labeled probe. Hybridization specifically tags the sequence of interest, even though it may constitute only a minute fraction of the nucleic acids on the filter; this is the basis of the exquisite selectivity of the method. Unbound probe is washed off, and the filter is exposed to x-ray film. The position of a DNA or RNA fragment complementary to the probe appears as a band on the film. This procedure is termed *Southern blotting* when DNA is transferred to nitrocellulose, *Northern blotting* when RNA is transferred, and *Western blotting* when protein is transferred from an SDS-polyacrylamide gel. In Western blotting, the protein of interest is visualized using an antibody that specifically recognizes it within the background of other cellular proteins.

Recently, Southern blotting has been used in conjunction with electrophoretic separation of very large DNA molecules to prepare restriction maps over distances of hundreds of kilobases (discussed in Chapter 29). In this way genes can be ordered along chromosomes, revealing the clustering of genes into functionally related groups (Chapter 8). As we shall see, the Southern blotting technique has been important for our understanding of a variety of biological processes such as RNA splicing (Chapter 8) and genomic

rearrangements to form antibodies and T cell receptors (Chapter 16) and in the detection of rearranged oncogenes (Chapter 19).

Northern blotting is a technique used to analyze RNA. Total cellular RNA, or poly(A) RNA, is separated by size on an agarose gel. The RNA molecules in the gel are transferred to nitrocellulose or nylon as described above and detected using an appropriate probe. Northern blotting is useful as an adjunct to cDNA cloning because the size of a specific mRNA can be compared with the size of cloned cDNAs, revealing whether the cloned cDNA is full-length. In addition, this simple procedure can indicate which tissues or cell types express a particular gene or the factors that regulate its expression. An example of this is the analysis of regulated genes. The isolation of cDNA clones from serum-regulated genes was previously discussed in the section on differential screening. The first step in characterizing these clones was Northern blotting. Cells were stimulated with serum, then total RNA was isolated at several time points. The RNA obtained at each point was analyzed by Northern blotting using individual cloned cDNAs as a probe (Figure 7-24). The results showed that each mRNA was present in low levels in untreated cells and rapidly accumulated following serum stimulation.

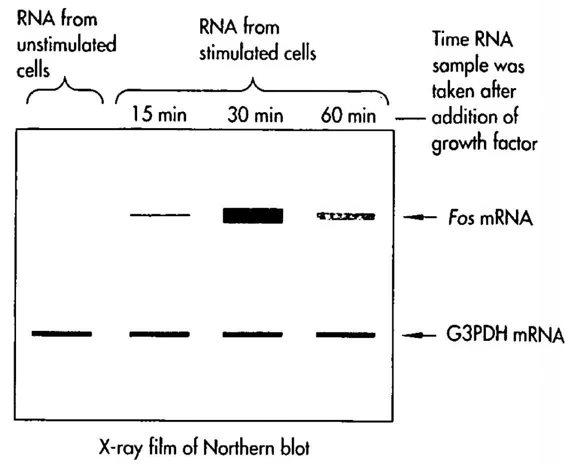


FIGURE 7-24
Northern blotting. This method is used to examine the size and expression pattern of specific mRNAs. In this experiment, RNA was isolated from cultures of resting fibroblasts stimulated with serum growth factors. Equal amounts of RNA from cells stimulated for the indicated times were applied to an agarose gel and transferred to a nitrocellulose filter by blotting. The filter was hybridized to two radiolabeled probes, one complementary to c-*FOS* (see Chapter 18) known to be regulated by growth factors in serum and one complementary to a nonregulated gene encoding glyceraldehyde 3-phosphate dehydrogenase (G3PDH). The 2.2-kb c-*FOS* mRNA was not found in resting cultures but appeared rapidly and transiently following stimulation. In contrast, the G3PDH mRNA was present at the same level in all samples.

Reading List

General

Gilbert, W., and L. Villa-Komaroff. "Useful proteins from recombinant bacteria." *Sci. Am.*, 242(4): 74–94 (1980). [Review]

Schmidtke, J., and D. N. Cooper. "A comprehensive list of cloned human DNA sequences." *Nuc. Acids Res. Suppl.*, vol. 19: 2111–2126 (1991).

Sambrook, J., E. F. Fritsch, and T. Maniatis. *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y., 1989.

Glover, D. M., ed. *DNA Cloning: A Practical Approach.* IRL, Oxford, Eng. (A series of volumes.)

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, K. Struhl, eds. *Current Protocols in Molecular Cloning.* Greene Publishing and Wiley Interscience, New York, 1989.

Methods Enzymology, vols. 68, 100, 101, 152, 153, 154, 155. Academic, New York.

Original Research Papers

PLASMID AND PHAGE VECTORS

Blattner, F. R., B. G. Williams, A. E. Blechl, K. Denniston-Thompson, H. E. Faber, L.-A. Furlong, D. J. Grunwald, D. O. Keifer, D. D. Moore, J. W. Schumm, E. L. Sheldon, and O. Smithies. "Charon phages: safer derivatives of bacteriophage lambda for DNA cloning." *Science,* 196: 161–169 (1977).

Bolivar, F., R. L. Rodrigues, P. J. Greene, M. C. Betlach, H. L. Heyneker, H. W. Boyer, J. Crosa, and S. Falkow.